Evaluation des tests de la section de français faculté de pédagogie Université Hélouânà la lueur d'uncritère de qualité de la feuille d'examen

Dr. Rania Mohamed Hamed Ismail

Résumé :

L'évaluation des tests est un point de départ pour souligner beaucoup de problèmes et prendre des décisions convenables aux niveaux de l'enseignement et de l'apprentissage à l'université.Pour cette raison, la recherche vise àélaborer un critère de qualité de la feuille d'examen écrit en françaisetévaluer les tests élaborés par les enseignants universitaires de la section de français faculté de pédagogie université de Hélouânà la lueur de ce critère.La chercheuse a élaboré un critère qui comprend deux axes :les normes formelles et techniquesde la feuille du test & les normes des typesde questions. L'échantillon était 45 tests de la section de français de l'année universitaire 2011/2012.Les résultats de cette recherche mettent en relief :Le manque de connaissances et de performances dansles normes formelles et techniquesde la feuille du test et les normes des types de questions chez les enseignants universitaires.La concentration sur les items à réponse élaborée et la question combinée.

Introduction

En cherchant à assurer la qualité dans nos universités, un système efficace d'évaluation d'apprentissage apparait comme un besoin urgent.

L'évaluation influence directement l'apprentissage quand elle y fournit le feedback nécessaire pour réaliser un apprentissage efficace et influence indirectement l'apprentissage car l'enseignement vise vers ce qui va être évalué.(El Assar, Safaa & autres, 2000, 28, p.32)

De même, l'évaluation d'apprentissage informe sur l'efficacité de l'acte d'enseignement.

L'évaluation d'apprentissage se varie entre les tests d'achèvement, les tests oraux, les tests de performance, les fiches d'observation, les fiches d'autoévaluation …etc. A la faculté de pédagogie, l'instrumentd'évaluation le plus courant est le test aux questions ouvertes .A la section de français on recourt, à côté des tests d'achèvement, aux tests de performance (de compréhension écrite).

En 2001, Gilles a préciséles problèmes liés aux examens écrits ayant recours aux questions à réponses ouvertesdans les points suivants :

1. ***Le manque de concordance intra et inter-correcteurs dans la correction des réponses ouvertes*** :

On distingue trois catégories de biais liés à la correction avec deux effets dans chaque catégorie.

1-Biais dus au seul correcteur :

Effet de sévérité :Sévérité systématiquement plus élevée ou au contraire moins élevée chez certainscorrecteurs.

Effet de tendancecentrale :Evitement desnotes extrêmes etconcentration desscores au milieude l'échelle.

2-Biais dus aux interactions professeur-élève :

Effet de halo :Des caractéristiques de l'étudiant influencent sa note (aspect physique,présentation, ...)

Effet deStéréotypie : Tendance à attribuer à un examiné les notes que celui-ci a acquiseantérieurement.

3-Biais dus aux séries de copies :

Effet de séquence : La copie qui suit une copie brillante risque d'être désavantagée, etinversement.

Effet de relativisation :Parmi toutes les copies jugées moyennes quelques mois auparavant le correcteur distinguera des faibles et des bonnes.

2. ***Le manque de validité*** :

Lorsque le nombre d'étudiants est élevél'examinateur est contraint d'utiliser un nombre réduit de questions ouvertes. Ce que pose à leur tour deux problèmes : l'absence d'une couverture large de tous les points importants du cours et la simple restitution de faits abordés dans le cours malgré la possibilité d'évaluer des processus mentauxplus complexes.

3. ***Le manque de sensibilité des mesures*** :

Rarement on demande aux étudiants d'exprimer systématiquement leurs certitudes à propos desréponses qu'ils fournissent lors d'un examen.

4. ***Le manque de diagnosticité*** :

Les examinateurs se préoccupent en général très peu de renvoyer après l'épreuve un feedbackdétaillé et individualisé vers les examinés de manière à permettre à ces derniers d'effectuer un bilan précis de leurs compétences.

Cependant, dans le contexte universitaire les étudiants qui subissent un échec dans un cours en1ère session disposent en général d'une deuxième chance sous la forme d'une seconde épreuve en 2ème session (souvent du même type que la première) et de quelques mois pour améliorer leurs connaissances.

5. ***Le manque d'équité des épreuvestraditionnelles*** :
Il faut reconnaître que le facteur chance peut en effet jouer un rôle importantdans la réussite lorsque seulement deux ou trois questions sont posées lors d'une épreuve.

En corrigeant de copies d'épreuves écrites: une copie de qualité moyenne sera surévaluée après une copie médiocre ou sous-évaluée après une copieexcellente.(**Gilles, Jean-Luc,2002,11, Pp.12-15**)

Pour des tels problèmes,des études soulignent l'importance d'entraîner l'enseignant universitaire à améliorer ses pratiques d'évaluation pour mesurer l'atteinte des objectifs d'apprenti-ssage. D'après Brassard, enseigner à l'université impliqued 'élaborer et de structurer des activités d'évaluation, de préparer, administrer et évaluer les apprentissages en fonction de l'atteinte des objectifs de départ. (Brassard, Nancy, 2012,3, p.5)

Dans son étude, Romainville a proposé de former les enseignants universitaires à l'évaluation à travers des ateliers deréflexion sur leurs propres pratiques d'évaluation des acquis (Romainville, Marc, 2002, 22, p.46)

Donc, l'évaluation des tests est un point de départ pour souligner beaucoup de problèmes et prendre des décisions convenablesaux niveaux de l'enseignement et de l'apprentissage à l'université.

<u>La problématique de la recherche</u> :
La chercheuse a constaté le problème par :

1. **<u>Expérience personnelle :</u>**
Aux travauxdes examens, la chercheuse a fait faceaux problèmes tels :

- Le manque de quelques directives générales du test comme l'année universitaire, le semestre et la classe.
- Parfois, les professeurs oublient de citer que le test est en plus d'un page et d'autres oublient de signer le test.

2. **Etude exploratoire :**

Dans une rencontre ouverte avec les étudiants de la 2ème année section de français à la faculté de pédagogie de Hélouân, on a posé cette question :Quels problèmes trouvez- vous dans la feuille d'examen ?

Les étudiants se sontplaints des problèmes tels :

- l'illisibilité du test(des petits graphiques).
- Ne pas comprendre une question,la question est hors le contenu du programme.
- Quelques étudiants n'ont pas répondu qu'à la moitié du test de l'année précédente parceque le professeurn'a pascité pas que le test est en deux pages.
- Le test se compose d'une seule question.
- Les consignes ne sont pas claires et ne précisent pas exactement ce que les étudiants doivent faire.
- Le niveau de la langue est élevé.
- Nepas s'entraîner aux types de questions de l'examen.

Tous ces problèmes influencent les résultats des étudiants qui à leur tour influence la qualité de l'enseignement et de l'évaluation à la faculté. Donc, on a besoin d'évaluer les tests élaborés par les enseignants universitaires à la lueurd'un critèrede qualité de la feuille d'examen, pour préciser les points forts et les points faibles dans ces tests et leur fournir du feed back.

Les questions de la recherche :

Pour surmonter le problème ci-dessus, on devrait répondre aux questions suivantes:

1. Quels sont les normes nécessaires pour élaborer un test écrit en français ?
2. A quel point les tests élaborés par les enseignants universitaires reflètent–ils ces normes?

Les limites de la recherche :

La recherche se limite:

1. A tous les tests de la section de français faculté de pédagogie qui sontécrits en français.
2. Aux deux semestres de l'année universitaire 2011/2012.

3. Aux matières académiques et pédagogiques suivantes : civilisation, phonétiques, expression écrite, méthodologie, essai, traduction, étude de textes (prose, poésie, théâtre), littérature, compétences d'enseignement, grammaire, lecture et écriture, lecture et conversation , linguistique, culture générale ,emploi de l'informatique dans l'enseignement du français et français sur objectif spécifique .

Les objectifs de la recherche :

Cette recherche vise à :

1. Elaborer un critère de qualité de la feuille d'examen écrit en français.
2. Evaluer les tests élaborés par les enseignants universitairesde la section de français faculté de pédagogie université de Hélouânà la lueur de ce critère.

L'importance de la recherche :

Cette recherche peut participer à :

1. Présenter un critère pour évaluer les tests écrits en français par les enseignants universitaires.
2. Attirer l'attention des responsables à la faculté de pédagogie sur l'importance d'évaluer les tests élaborés par les enseignants universitaires.
3. Aider les enseignants universitaires à évaluer leurs tests et préciser leurs points forts et leurs points faibles pour s'améliorer.
4. Proposer un critère qui peut aider les enseignants universitaires à élaborer les tests.

Les procédures de la recherche

Premièrement le cadre théorique :

1. **La chercheuse a consulté des ouvrages et des études antérieures qui portent sur les sujets suivants :**

Le test comme un instrument d'évaluation sommative,les qualités d'un bon test,les consignes du test ,le choix des types d'items, les types d'items du test et les contingences qui orientent les différentes pratiques d'examen à l'université .

Deuxièmement le cadre expérimental:

2. Elaborer un critère pour évaluer les tests écrits de la section de français faculté de pédagogie.Le critère comprend deux axes : Les normes formelles et techniquesde la feuille du test & Les normes des typesde questions.
3. Présenter le critère au jury puis faire les modifications nécessaires.
4. Analyser les tests selon le critère proposé.
5. Discuter les résultats.
6. Faire les recommandations et les suggestions convenables.

Les terminologies de la recherche

Evaluation :

- « Examiner le degré d'adéquation entre un ensemble d'informations et un ensemble de critères adéquats à l'objectif fixé en vue de prendre une décision. »(Raynal, Françoise &Rieunier, Alain, 1997, 21, p.138)
- « Evaluer ; signifie recueillir des informations puis les analyser pour une prise de décision. » (Centre National de DocumentationPédagogique, 2004, 9)
- « Uneévaluationest une opération qui consiste à recueillir des données et à poser un jugement sur une personne, un objet, un processus, une situation ou une organisation en comparant les caractéristiques observables à des normes établies, à partir de critères explicites en vue de fournir des informations utiles à la prise de décision dans la poursuite d'un but ou d'un objectif. » (Bureau d'évaluation de l'enseignement et des programmes d'études, 2007, 4)
- « L'évaluation est le processus de recueillir les informations et les utiliser pour poser des jugements qui à leurs tours aidentà prendre des décisions. »

(Tenbrink, Terry D., 1999, 25 , p. 311)

Définition opérationnelle

L'évaluation est un processus d'analyser les feuilles des tests par rapport à un critère pour poser un jugement sur la qualité des tests qui à son tour aide à prendre des décisions.

Test :

- « On appelle test un examen standardisé, fidèle, valide et étalonne. Il sert à évaluer le niveau de connaissance ou de

compétence acquises par un individu dans un domaine donné. Il peut être pronostique, diagnostique ou projectif et sert généralement à classer les individus les uns par rapport aux autres. »(Cuq, Jean-Pierre, 2003, 7, p.236)

- « Epreuve standardisée permettant de mesurer et d'évaluer une connaissance particulière ou une dimension de la personnalité .Cette évaluation s'effectue en comparant les résultats obtenus par un individu, à ceux obtenus par un groupe témoin. » (Raynal, Françoise &Rieunier, Alain, 1997, 21, p.360)
- « Une situation standardisée permettant d'évaluer le comportement d'un individu en comparaison avec celui d'individusplacés dans la même situation » (Beaudet, Suzanne &Monnerie-Goarin, Annie 1993, 2, p.59)

Définition opérationnelle

Un instrument élaboré par l'enseignant universitaire pour mesurer le niveau d'atteinte des objectifspar lesétudiants.

L'étude théorique

1. Le test comme un instrument d'évaluationsommative :

D'après Pendanx, l'évaluation sommative remplit deux fonctions principales : d'unepart, une fonction sociale, puisqu'elle donne lieu aux certifications des différents niveaux de compétence et de maîtrise de savoir-faire.

D'autre part, une fonction pédagogique, car elle fournit à l'enseignant comme à l'apprenant, des indications sur le processus d'apprentissage.(Pendanx, Michèle, 1999, 18, p.49)

Portée sur le passé, elle se situe à la fin d'une période donnée d'apprentissage (un semestre, une année).Contrairement à l'évaluation diagnostique qui ne doit pas figurer dans la note de l'apprenant et l'évaluation formative qui ne se traduit pas systématiquement en note mais plutôt en appréciation et commentaires informatifs (Clément, Marlène, 2004,6),l'évaluation sommative fait donc le bilan d'une formation et les résultats obtenus, traduits en notes, certifient l'état de cette formation.

C'est pourquoi, on caractérise souvent cette évaluationde certificativeou de normative car les résultats permettent de

situer un apprenant par rapport à une norme et/ou en vue d'un classement. (Cuq, Jean-Pierre &Gruca, Isabelle, 2005, 8, p.210)

Il est à noter que l'évaluation sommative apparaît sous plusieurs titres : certificats, diplôme, crédits ....

Henri & Cormier proposent des balises générales à respecter pour rendre l'évaluation sommative plus efficace :

1. L'évaluation sommative doit traduire un jugement porté à la fin de l'apprentissage et ne peut être constituée d'évaluations formatives cumulées.
2. L'évaluation sommative doit prendre en compte les performances de l'apprenant telles quelles se manifestent à la fin de l'année.
3. Le jugement de maîtrise porté par l'enseignant lors de l'évaluation sommative doit s'appuyer sur des mesures effectuées à l'aide d'instruments formels : tests, grilles d'observation, critères de performance, etc.
4. L'évaluation sommative doit s'appuyer sur les éléments essentiels et mesurables du programme.
5. L'évaluation sommative doit respecter les principes d'évaluation particuliers au programme concerné, s'il y en a.
6. Les situations d'évaluation doivent être compatibles avec les habiletés évaluées : les habiletés pratiques doivent être évaluées par une épreuve pratique, par exemple.
7. Les tâches de l'évaluation sommative doivent être, pour l'apprenant, d'un type familier et faire appel exclusivement à des habiletés et des connaissances sollicitées au cours de l'apprentissage. (Henry, jacques & Cormier, Jocelyne,12)

2. Les qualités d'un bon test :

Le test comme l'instrument le plus commun dans l'évaluation sommative aux universités pour qu'il réalise son but, on doit couvrir les qualités suivantes :

- La validité : un test est valide ou pertinents'il mesure réellement ce qu'il est censé mesurer. Il doit refléter les objectifs généraux et particuliers du programme.

- <u>L'objectivité</u> : un test est objectif si des corrections par des correcteurs différents ou deux corrections par le même correcteur à quelques jours ou semaines d'intervalle, donnent des résultats identiques.
  Le garant de l'objectivitéd'un test est soit sa nature, soit la rigueur apportée dans la correction grâce à une grille de correction. Le test s'accompagne d'un corrigé et d'un barème précis.
- <u>La fiabilité</u> : la fiabilité d'un test est démontrée lorsqu'on obtient les mêmes résultats en soumettant plusieurs fois les mêmes sujets aux mêmes tests et dans les mêmes conditions.
  (Quinton, André, 2005, 20)
- <u>L'économie</u> : un test doit être peu coûteux surtout en temps et en énergie. On peut distinguer l'économie de construction, de passation et de correction. (Beaudet, Suzanne &Monnerie-Goarin, Annie 1993, 2, p.60)
- <u>La qualité matérielle</u>: Assurez-vous de la qualité des photocopies ou des enregistrements.

Le Système Méthodologique d'Aide à la Réalisation de Tests (SMART) a annoncé que lors de la construction d'un test standardisé universitaire dequalité il s'agit d'offrir des garanties en ce qui concerne : Lavalidité, la fidélité

- <u>La sensibilité</u> : la mesure doit être précise ;
- <u>Ladiagnosticité</u> : le diagnostic précis des difficultés d'apprentissage, des processus maîtrisés et de ceuxqui ne le sont pas doit être possible ;
- <u>La praticabilité</u> : la faisabilité en termes de temps, de ressources en personnel et en matériel doit être assurée ;
- <u>L'équité</u> : tous les étudiants doivent être traités de façon juste, en principe de la même manière (standardisation) ;
- <u>La communicabilité</u> : les informationsnon confidentie-lles relatives au déroulement du processus doiventêtre communiquées et comprises par les partenaires, enseignants, étudiants, équipe de soutiendocimologique (par exemple le SMART), …, engagés dans la réalisation des épreuves. **(Gilles, Jean-Luc, 2002, 11, Pp .21-22)**

D'après l'université catholique de Louvain pour concevoir un bon test écrit il faut suivre les instructions suivantes:

1. Formuler les questions de manière claire et non ambiguë afin que chaque étudiant soit en mesure de comprendre ce qu'on demande.
2. Mesurer des connaissances et des compétences qui sont en rapport direct avec les objectifs du cours ou qui font partie des prérequis.
3. L'ampleur et le contenu du test doivent être compatibles avec la durée prévue à l'horaire.
4. L'ensemble des questions doivent couvrir tous les types de compétences qu'on souhaite tester.
5. Mettre en œuvre des niveaux plus élevés de réflexions et mieux cerner les objectifs d'apprentissage du cours.
6. La couverture de la matière : on a parfois inconscie-mment tendance à privilégier des sujets qui ont été abordés soit en début soit en fin de quadrimestre.
7. Valider le test : on considère que vous devriez pouvoir y répondre en y consacrant un temps qui ne devrait pas dépasser le quart du temps dont disposeront les étudiants.
8. Commencer le test par des questions qui donnent confiance aux étudiants (les premières questions sont d'un niveau de difficulté raisonnable).
(L'université catholique de Louvain, 2007, 14 , Pp .29-31)

Enfin, on peut exposer le système de gestion et de contrôle de la qualité des tests standardisés élaboré parl'équipe du SMART qui a invité les enseignants à suivre une démarche structurée en huitétapes :

1. Analyse des objectifs de l'enseignement qui aboutit à une table des spécifications des matières et processus mentaux à évaluer;

2. Mise en forme de l'épreuve et définition des modalités de questionnement;

3. Construction de l'épreuve et rédaction des questions

4. Entraînement des étudiants aux procédures d'évaluation;

5. Préparation matérielle et mise en œuvre de l'examen;

6. Traitements de correction suivis d'une discussion des résultats avec rectifications éventuelles;

7. Information des étudiants à propos de leurs performancesà l'aide de feedbacks individualisés;

8. Macro régulation du cycle de réalisation de l'épreuve à l'aide des avis des étudiants.(Gilles, Jean-Luc,2004,10 ,p. 106)

3. Les consignes du test

D'après le Cadre Européen, les consignes sont les instructions données aux étudiants pour réagir devant un support donné. Ces instructions doivent préciser où et comment la réponse sera enregistrée.(Milanovic, M., 2002, 15)

La consigne définit la tâche, les moyens et le degré de réussite ou d'achèvement(Oasisfle.com,2013, 17)

Selon Puren, La consigne est l'énoncé du concepteur de manuel ou de l'enseignant présentant l'exercice ou la tâche à réaliser. Certains consignes détaillées vont jusqu'à inclure un guidage.(Puren,Christian&autres, 1998, 19, p.196)

Donc, ces définitions soulignent l'importance des consignes et précisent les détails qu'elles doivent couvrir.

Les qualités d'une bonne consigne :

1. *Claire*, elle ne doit y avoir ni ambiguïté ni besoin de clarification. La compréhension d'une consigne dépend aussi beaucoup du vocabulaire employé. Il semble nécessaire de remplacer les verbes mentalistes, qui ne sont pas assez précis, par des verbes exprimant des comportements mieux cernés. « Comparer » pourra devenir « noter par écrit, sous la forme d'un tableau, les éléments identiques et les éléments différents contenus dans….. » De plus, certains mots sont porteurs de plusieurs sens et peuvent même avoir une signification tout à fait différente.
2. *Précise*, tout en restant ouverte, de façon à ne pas communiquer le résultat.
3. *Restrictive*, dans le sens où elle doit mobiliser des connaissances bien précises chez les étudiants et focalise leur attention sur ce qui est important en fonction de l'objectif poursuivi.

Elle peut être courte ou longue.(Oasisfle.com, 2013, 17) (Salabura, Sébastien, 2002, 23, p.13)

Les types des consignes :

Mérieux distingue quatre types de consignes :

- Les consignes-critères : correspondent à des critères d'évaluation ou à des critères de réussite d'un travail.
- Les consignes-buts : correspondent aux objectifs d'un ou de plusieurs travaux .Elles sont généralement émises au début d'une séance ou de séquence.
- Les consignes-guidages : permettent d'attirer l'attention de l'apprenant sur un point particulier ou sur une consigne de sécurité .Elles sont régulièrement utilisées aux cours des travaux pratiques.
- Les consignes-procédures : correspondent aux étapes obligatoires ou possibles pour parvenir aux résultats escomptés. (Salabura, Sébastien, 2002, 23, Pp.11, 12)

Comment l'étudiant réagit avec les consignes ?

Ilya trois types de processus cognitifs dans l'utilisation des consignes ; lecture-compréhension, planification puis exécution des tâches.

A cette 1ère étape de processus cognitif est liée des problèmes tels :

Eluder la1ère étape (Chaplier, Claire, 2010 , 5)ou commettre des erreurs relevant de la compréhension des consignes de travail.

Les remédiations possibles à ces problèmes sont:

- Fournir aux étudiants le temps pour bien lire les consignes.
- Suivre les qualités d'une bonne consigne, en rédigeant le test.

4. Le choix des types d'items

Savoir lequel d'items est le mieux adapté pour tester une connaissance ou une compétence donnée dans un test est une décision prise au moment de la planification du test.

On peut classer les différents types d'items selon la correction en deux catégories :

- Items à correction objective : aucun jugement n'est nécessaire pour les corriger, ils contrôlent la compréhension.

- Items à correction subjective : demandent une réponse élaborée et des méthodes subjectives de correction, ils contrôlent la production.(Milanovic, M., 2002, 15)

Selon Brown &Hudson, les types d'items peuvent êtreclassés selon la réponse en trois catégories:

- La réponse sélectionnée : l'apprenant choisit une réponse(le choix multiples, l'alternative, l'appariement ....)
- La réponse élaborée : L'apprenant doit fournir la réponse (réponse courte, performance....)
- La réponse personnelle (portfolio, conférence, journal du bord.....)

  (Shaaban, Kassim, 2001, 24 , p.18)

Même si certains types d'items sont plus utilisés que d'autres, il ne faut pas croire qu'ils sont meilleurs pour autant.

SMART (Le Système Méthodologique d'Aide à la Réalisation de Tests) à l'Université de Liègesoulignela nécessité d'une complémentarité entre les épreuves standardisées ayant recours aux QCM etquestions à réponses ouvertes moyennes ou longues (QROM/QROL) écrites améliorées à l'aide de techniques telles que les échelles descriptives lorsque les objectifs de l'évaluation l'exigent.

Ainsi, des QROM / QROL corrigées avec la technique des échelles descriptives permettent de mesurer des performances complexes tandis que le nombre élevé de QCM(30 items) permet une couverture large de la matière. Une partie du gain de temps considérable lié à la correction automatisée des QCM peut ainsi être réinvestie dans la correction plus poussée d'une ou deux QROM/QROL avec échelles descriptives.(Gilles, Jean-Luc,2002, 11 ,p.20)

5. Les types d'items

Dans cette étude, on va adopter la classification qui répartit les types d'itemsen deux catégories selon la réponse :

1. *Items à réponse sélectionnée*

1-1Items à choix multiples

Le tronc se présente généralement sous la forme d'une question ou d'un début de phrase suivie d'une série de réponses possibles ou de propositions de fin de phrase dont l'une est la

bonne. Les réponses incorrectes s'appellent des distracteurs.Ces réponses peuvent se réduire à un mot ou à une proposition ouêtre plus développées sous la forme de paragraphes ou de mini-textes.

L'apprenant répond en choisissant la réponse correcte (parfois deux réponses correctes) parmi trois, quatre ou cinq propositions ; il s'agit soit de cocher une case, soit de souligner ou d'entourer la réponse retenue. La bonne réponse doit se trouver n'importe où.

Ces items testent la compréhension, et les connaissances et l'application sur un sujet donné (grammaire, vocabulaire, phonétique……….) .Il est surtout pratique dans les cas des définitions, dedifférences, de discrimination, de généralisation, de relations de cause à effet.

Les items à choix multiples aident à balayer tout le champ d'un enseignement.Leur correction est fiable objective et rapides mais son élaboration nécessite beaucoup d'attention et de temps. Mais il est à souligner qu'ils ont plus de valeur si on les utilise conjointement avec d'autres types d'items. (Quinton, André, 2005, 20). (Cuq, Jean-Pierre &Gruca, Isabelle, 2005, 8, p.347)

(Ministère de l'Éducation du Saskatchewan, 1993, 16)

1-2<u>Items de type vrai ou faux</u>

L'item d'alternative comprend deux réponses possibles entre lesquelles l'apprenant doit choisir (oui/non, vrai/faux, correct/incorrect, d'accord/pas d'accord).On compense en fournissant à l'élève une troisième possibilité de réponse, en ajoutant (non mentionné) ou (on ne sait pas).

Ces items testent la compréhension et la connaissance qu'un apprenant a du contenu de la matière.

Ce type est efficace avec des documents longs, car il permet de soutenir l'attention et favorise l'accès à la compréhension si les textes sont difficiles.

Il est rapide à corriger, assez rapide à construire mais il est possible pour les étudiants de deviner ou déduire la bonne réponse.

(Lussier, Denise, 1992, 13, p.59) (Ministère de l'Éducation du Saskatchewan, 1993, 16)

1-3 Items de type appariement :

Aussi appelés pairage ou association, ils consistent en un ensemble de problèmes ou de questions (les prémisses) alignés dans une colonne et un ensemble de réponses alignées dans une autre colonne. Faire associer un texte à un dessin, un schéma, une photo……….)

Ils permettent de vérifier la capacité d'établir des liensdes faits connexes, des associations et des corrélations.Les items de type appariement sont d'autant plus efficaces qu'ils sont utilisés parallèlement à d'autres types d'items.(Cuq, Jean-Pierre &Gruca, Isabelle, 2005, 8 , p.350) (Lussier, Denise, 1992, 13 , p.59) (Ministère de l'Éducation du Saskatchewan, 1993, 16)

1-4Items de type réarrangement

Aussi appelés les puzzles ou la reconstitution de texte. Ils consistent à replacer dans un ordre logique ou chronologique des éléments d'une phrase ou d'un paragraphe ou les différentes parties d'un discours ou les lettres d'un mot.

Le réarrangement mesure l'ordre logique des idées (la cohérence) et l'utilisation adéquate des mots charnières (la cohésion) .Il s'applique particulièrement aux textes narratifs et aux textes argumentatifs. Si l'on travaille sur le mot ou la phrase, c'est essentiellement l'orthographe ou la syntaxe qui seront sollicités. On peut le complexifier soit en introduisant simple un ou des distracteurs (phrase intruse), soit en proposant un puzzle composé de deux textes. La difficulté de cet exercice relève du nombre des pièces découpées et de la règle de décomposition.(Cuq, Jean-Pierre &Gruca, Isabelle, 2005, 8, Pp.348, 349) (Lussier, Denise, 1992, 13 , p.59)

2. Items à réponse construite

2-1Items à réponse courte

Ce type exige de l'apprenant de donner une réponse à une question en particulier. La réponse peut aller d'un mot à un paragraphe.Ces items mesurent la connaissance ou des processus mentaux de niveau élevé.

Les types des items à réponse courte :

1-Les items à choix multiples et les items de type vrai ou faux peuvent être transformés à items à réponse courte en demandant de justifier la réponse.

2-Les questions orientées ou guidées : invitent l'apprenant à montrer le bien- fondé ou non d'une ou plusieurs informations. Il s'agit donc d'une activité de repérage et de justification ou d'élucidation. Elles se présentent sous la forme « Montrez que………. »

3-Les textes lacunaires :

- **Le texte à trous**

Fabriqué à partir de phrases isolées ou à partir d'un texte dans lequel on a aménagé des trous qu'il faut remplir. Il consiste à effacer les mots d'un même champ lexical ou d'une même catégorie grammaticale, il est possible de supprimer seulement des lettres ou les accents ou la ponctuation ou des phrases complètes. On peut donner en marge les éléments manquants avec/ ou sans distracteur.

- **Le test de closure**

Consiste à supprimer de manière aléatoire, un mot sur cinq(ou à intervalles réguliers entre cinq et dix) dans un texte généralement authentique, à la place des mots manquant, qui n'appartiennent pas à une catégorie précise, on ménage des espaces vides.Il est possible de présenter sous forme de liste et dans le désordre les mots manquants.Si l'on désire complexifier l'exercice, on peut introduire dans la liste des mots intrus, à condition bien évidemment de la signaler dans la consigne.(Cuq, Jean-Pierre &Gruca, Isabelle, 2005, 8, Pp.347, 348)(Ministère de l'Éducation du Saskatchewan, 1993, 16)

2-2Les activités d'analyse et de synthèse

Ces activités engagent non seulement la compréhension des documents et l'expression écrite, mais également la connaissance des règles méthodologiques (la consigne indique la qualité du texte à produire).

Le résumé exige le respect de l'organisation du texte initial et il interdit la mention du document source alors que le compte rendu les tolère ; quant à la synthèse elle porte sur plusieurs documents de nature diverse (textes, schémas, photos, etc.), mais présentant la même problématique : sous la forme d'un texte unique et cohérent comportant une brève introduction et

une conclusion concise, elle appelle une confrontation contractéedes contenus , l'élaboration d'un plan personnel et l'indication des documents sources autorisent la comparaisondes informations (points communs, différences, etc.)(Cuq, Jean-Pierre &Gruca, Isabelle, 2005, 8, Pp.350, 351)

2-3Les activités de réécriture

2-3-1 Les modifications de texte par contractionou par expansion :

Les exercices de réduction : consistent à réduire un texte à des degrés divers. Il s'agit d'une part de conserver la structure générale du texte, son contenu essentiel et ses principaux ingrédients phrastiques, et, d'autre part, de supprimer un certain nombre d'éléments qui ne paraissent pas importants. Les exercices d'amplification : consistent à donner du volume à un texte sans autant le dénaturer. Dans cette perspective, trois procédés importants permettent d'étoffer un texte : 1- la démultiplication des mots clés.2- Introduire des péripéties qui retardent le déroulement du texte3-insérer une phrase supplémentaire entre deux phrases sans briser la cohérence du texte.

1-3-2 Les activités de réparation du texte :

Soit par simplification : un texte est jugé comme trop difficile et on le simplifie en décomposant syntaxiquement les phrases ou en les paraphrasant en séquences plus courtes.

Soit par complexification : on propose des phrases simples et on demande de construire des phrases complexes qu'il faut intégrer afin d'obtenir un paragraphe ou un texte (on retrouve cet exercice lors de l'apprentissage grammatical pour le passif et la nominalisation ou pour l'emploi des connecteurs)

Ces activités ont l'avantage de mettre en relief les influences des règles textuelles sur la phrase.

2 -3- 3Les activités de remise en discours ou en texte

Contraintes de genre de texte : par exemple, on propose un extrait narratif et on demande de le réécrire sous forme de dialogue ou de scénario filmique.

Changement de point de vue : consiste à faire réécrire un texte en changeant la perspective du narrateur ou en faisant réécrire le texte du point de vue d'un autre personnage ou par un narrateur omniprésent.(Cuq, Jean-Pierre &Gruca, Isabelle, 2005, 8, Pp.354-356)

2.4Items à réponse élaborée

L'apprenant répond longuement par écrit à une question posée. Ce type d'item sert à évaluer le sens de l'argumentation, de l'évaluation et de la synthèse, exposer les convictions et les points de vue et communiquer par écrit, s'exprimer de façon personnelle et unique. (Décrire, comparer, expliquer, discuter, développer, résumer, évaluer………)

Ce type exige des efforts de réflexion et demande de l'apprenant d'ordonner ses idées, faire un plan, développer des arguments et de faire preuve de sa maîtrise de langue. Selon le niveau de réflexion on peut distinguer les questions où :

-répondre nécessite *une réflexion originale.*

-répondre nécessite *une réflexion originale mais suitun plan convenu*. Sous ce niveau, on peut poser les tests de l'expression écrite et parmi leurs types d'activités :

Les exercices de réparationde texte : consistent à inventer le début ou la fin ou le cœur d'un texte. Cet exercice, étroitement lié à la compréhension du texte support, accorde une place importante à l'imaginaire de l'apprenant qui est cependant contrôlé ou guidé par des contraintes de cohérence textuelle.

Les matrices de textes : production d'un texte à partir de contraintes clairement énoncées et /ou d'un canevas plus ou moins conséquent qui répond à des règles textuelles précises. Les matrices peuvent être simples et élémentaires .Par exemple : proposer de rédiger un texte cohérent d'une longueur précise, avec un vocabulaire ou des mots grammaticaux imposés.

Les textes peuvent être guidés par un autre support (image, dessins, bandes dessinées…) servir divers objectifs. (Cuq, Jean-Pierre &Gruca, Isabelle, 2005, 8, Pp.352-354)

-répondre se limite à *restituer une partie de cours* sans faire appel à la réflexion.

(Quinton, André, 2005, 20) (Ministère de l'Éducation du Saskatchewan, 1993, 16) (Lussier, Denise, 1992, 13, p.62)

6. Les contingences qui orientent les différentes pratiques d'examen àl'université :

Le Groupe de Recherche Interdisciplinaire en Pédagogie Universitaire (GRIPU) de l'Université deMontréal adonné quelques pistesqui peuvent expliquer les différentes modalités d'examens utilisées par les professeurs :

a) La culture d'évaluation

Dans chaque université ou peut être faculté existe des pratiques d'évaluation qui dépendent d'une culture.Ainsi, on peutse demander ce qui arrive lorsqu'un professeur veut trop s'éloigner de cette culture et introduire des pratiques différentes ? Sera-t-il encouragé ou ostracisé ? Tout dépend probablement des nouveautés introduites, de la distance entre ce qui est nouveau et ce qui est 'traditionnel'.

Les années deformation des étudiants les mettent en présence d'une culture de l'évaluation propre à leur domaine. Pour celui qui fait desétudes de premier cycle dans un domaine et qui revient y enseigner après des études graduées, il sera naturel dans un premier temps de reproduire ce que les autres font autour de lui et ce que lui-même a vécu lors de ses études, une forme d'autocontrainteinconsciente.

D'après Samuelowicz& Bain (2002), l'enseignant universitaire organise l'évaluation des acquis de ses étudiants selon son profil personnel d'évaluateur. Il propose trois profils d'évaluateur.

- Dans le premier, le professeur cherche surtout à *mesurer la capacité de l'étudiant à reproduire de l'information*, telle qu'elle a été présentée au cours et/ou dans le polycopié. L'évaluation est considérée comme un moyen de pression pour faire travailler les étudiants. La différence entre les étudiants qui réussissent et ceux qui échouent tient dans la quantité de matières qu'ils sont capables de restituercorrectement.
- Le second profil rassemble les examinateurs qui cherchent à mesurer la *capacité de l'étudiant* à

reproduire de l'information et à *l'appliquer à des situations nouvelles*. La compréhension implique donc, de leur point de vue, une transformation ou une réinterprétation des connaissances acquises de manière à répondre aux exigences d'une tâche nouvelle.

- Au sein du troisième profil, les examinateurs conçoivent des épreuves qui mesurent la *capacité de l'étudiant à intégrer, transformer et utiliser de manière personnelle des connaissances.*(Romainville, Marc, 2002, 22, p.21)

b) <u>La taille du groupe</u>

Il est clair que la taille du groupe exerce une pression sur les pratiques d'évaluation. Dès lors, il estlogique que l'enseignant confronté à de grands groupes d'étudiants, aura tendance à s'orienter vers des formes de questionnement dont la correction ne luiprendra pas trop de temps (des questions fermées de type QCM ou vrai-faux).

Cette contrainte ne touche pas seulement la façond'évaluer mais également ce qui est évalué en termes d'habiletés sollicitéeset développéesetl'orientation des stratégies d'apprentissage par les étudiants.

c) <u>Le règlement pédagogique</u>

Des règlements pédagogiques existent dans chaque institution universitaire, ces règlementscontiennent des contraintes en matière d'organisation des examens auxquellesles enseignants doivent seplier.

d) <u>La contrainte des disciplines</u>

L'emploi de tels types de questions ou d'autres remplissent parfois la fonction pour laquelle ils sont utilisés. Par exemple, on emploi les QCM si de bien couvrir la matière et se corrigent rapidement avec un minimum de subjectivité. Elles font plus appel à la mémoire qu'à la capacité de raisonnement.

e) <u>La lutte contre une dissonance pédagogico-docimologique</u>

On peut imaginer que des modalités d'examen soient plus en harmonie avec certaines formes depédagogie universitaire. Dans une approche participative, où lesétudiants travaillent très souvent en groupe sur des projets de développement, on

parie quel'évaluation sera teintée par ce choix d'approche pédagogique. Le contraire pourrait amener les étudiantsà contester la validité des modalités d'évaluation en regard des habiletés développées dans le cadre ducours. Ce problème semble aussi relever de la « culture d'évaluation ».(Gilles, Jean-Luc, 2002, 11, Pp.32-33)

L'outil de la recherche

Le questionnaire

Le but:

Ce questionnaire a pour but d'avoir un critère de la qualité de la feuille d'examen pour évaluer le test écrit en français par l'enseignant universitaire.

Les sources:

Les sources de ce critère sont :

1. L'étude théorique.
2. Quelques ouvrages dans ledomaine de l'évaluation.(AlyMaher,2007,29)(Sawsan Chaker,2007,27)(Aide-mémoire,2009,1) (Helmy El Wakil, 1999, 26)

La description:

Sous sa forme initiale, le questionnaire comprend deuxaxes principaux :

1. ***Les normes formelles et techniquesde la feuille du test*** **(21 normes):**
   - Les directives générales (6normes)
   - Les consignes (3 normes)
   - Des critères généraux (12 normes)
2. ***Les normes des typesde questions*** **(38 normes) :**
   - Items de type réarrangement (2 normes)
   - Items à choix multiples (11normes)
   - Items de type appariement (5normes)
   - Items de type vrai ou faux (7normes)
   - Textes lacunaires (6normes)
   - Question combinée (et compréhension écrite) (4 normes)
   - Items à réponse élaborée(et production écrite) (3normes)

Donc le questionnaire comprend 59normes.

On a élaboré deux exemplaires du questionnaire; un en français et l'autre en arabe pour permettre aux spécialistes dansle domaine de l'évaluation de juger le questionnaire.

La validité du critère:

Pour vérifier la validité du questionnaire, la chercheuse l'a soumis à un jury de spécialistes afin de donner leurs opinions en ce qui concerne les points suivants :

1. La suppression des normes non nécessaires.
2. L'addition des normes non inclues.
3. La modification des normes en cas de nécessité.

La modification du questionnaire

D'après les avis des spécialistes, la chercheuse a fait quelques modifications au contenu du questionnaire:

1. Ajouter « ou à travers une interrogative convenable » à la norme 1 dans les consignes pour devenir « Formuler les consignes à l'impératif ou au futur ou à travers uneinterrogative convenable »
2. Ajouter une norme aux critères généraux « s'assurer de la signature de l'évaluateur(le professeur de la matière) »
3. Ajouter « ou n'importe quelle phrase »à la norme 1 dans les items de type réarrangement pour devenir « Conserver le début et la fin du texte ou n'importe quelle phrase. »

Donc le questionnaire sous sa forme finale renferme 60 normes dont la chercheuse va profiter pour évaluer les tests écrits en français de la section de françaisfaculté de pédagogie.

Remarque :

La chercheuse, en élaborant ce critère, a conservé seulement les normes dont on peut évaluer les tests à la lumière .La chercheuse a pris en considération ce même principe en profitant des modifications du jury.

La stabilité (la fidélité)

Pour calculer la stabilité du critère, on a choisi un échantillon de 30 tests de l'année universitaire 2010/2011de la section de français .La chercheuse a appliqué la méthode de répétition de l'analyse des tests, trois semaines à peu près ont séparé les deux analyses (8/10/2011-30/10/2011)

On a calculé le facteur d'accord entre les deux analyses d'après la formule suivante : $CR = \frac{2M}{N1+N2} x100$

CR=Coefficient de relation

M= les points d'accord

N = Les items analysés

$CR = \frac{2 \times 55}{60+60} = \frac{108}{120} x100 = 91,7\ \%$

Ce qui éprouve la stabilité statistique du critère.

L'analyse et La discussion des résultats

La chercheuse a analysé les tests de l'échantillon et a calculé les résultats à travers l'équation suivante :

$\frac{\text{Nombre de fréquences de l'item}}{\text{Nombre d'items du questionnaire}} x100$

Discussion des résultats

La chercheuse a calculé le pourcentage d'après l'existence de chaque item dans les tests.

1. **Les directives générales** :

- La chercheuse a coché « Existe relativement » pour chaque test ne cite que deux ou trois de ces cinq points « La faculté, l'université, la section, la matière et la classe. »
- De même, pour chaque test ne cite de la date du test que le semestre et l'année universitaire.

2. **Les critères généraux :**

- « Numéroter les items en continu » pour cette norme ,3 tests de 45tests ne contiennent qu'une seule question donc la chercheuse a calculé le pourcentage selon 42 tests seulement ; de même pour la norme « Laisser d'espaces entre les questions. »

Le tableau (1)L'analyse des normes formelles et techniques de la feuille du test.

	Les normes	Existe(2)	Existe relativement (1)	N'existe pas(0)
A	Les directives générales doivent préciser :			
	1. La faculté, l'université, la section, la matière et la classe.	93,3	6,7	---
	2. La durée du test.	80	---	20
	3. La date du test.	11,1	77,8	11,1
	4. La totalité des notes.	48,9	---	51,1
	5. La note indiquée pour chaque question.	31,1	---	68,9
	6. Le nombre des questions à répondre.	20		80
B	Les consignes			
	1. Formuler les consignes à l'impératif ou au futur ou à travers une interrogative convenable.	93,3	---	6,7
	2. Utiliser des verbes exprimant des comportements mieux cernés.	84,4	---	15,6
	3. S'assurer que les consignes sont claires et contiennent toutes les données nécessaires.	37,8	60	2,2
C	Des critères généraux :			
	1. Numéroter les items en continu.	92,9	---	7,14
	2. Laisser d'espaces entre les questions.	100	---	---
	3. Utiliser des graphiques clairs et compréhensibles (lisibles).	97,8	2,2	---
	4. Citer si le test est en plus d'un page et les numéroter.	93,8	---	6,2
	5. Utiliser une phrase exprimant la fin du test.	48,9	---	51,1
	6. Utiliser des illustrations et des dessins de qualité.	100	---	---
	7. Utiliser une langue grammaticalement correcte et sans fautes d'orthographe.	95,6	4,4	---
	8. Citer les références des textes ou des supports.	33,3	---	66,7
	9. Mesurer des objectifsaux niveaux ultérieurs.	35,6	---	64,4
	10. Comprendre différents types d'items. (trois types au moins)	11,1	35,6	53,3
	11. S'assurer de la signature de l'évaluateur(le professeur de la matière).	37,8	---	62,2
	12. Eviter les questions en cascade.	97,6	2,4	---
	13. S'assurer que toutes les questions sont obligatoires.	71,1	17,8	11,1

- « Eviter les questions en cascade»pour cette normeun seul test a exposé des questions en cascade.
- « Utiliser des graphiques clairs et compréhensibles (lisibles) »,pour cette normeun seul test comprend un texte dont les graphiques sont très petits et illisibles.
- Pour la norme « Citer si le test est en plus d'un page et les numéroter »,16 tests de 45 tests d'un pourcentage 35,6 % sont en deux pages.
- « Utiliser des illustrations et des dessins de qualité », un seul test a utilisé une illustration et elle était vraiment de qualité.
- « Utiliser une langue grammaticalement correcte et sans fautes d'orthographe » 2 tests de 45 tests n'ont pas utilisés les lettres majuscules au début des questions.
- « Citer les références des textes ou des supports. » 20 tests de 45tests contiennent des textes et un test de 45 contient une illustration ça fait, 46,7% des tests.
- « Mesurer des objectifs aux niveaux ultérieurs » (16 tests) 35,6% des tests contiennent des questions qui mesurent des objectifs aux niveaux ultérieurs et pas tout le test ; (16 tests) 35,6 % de tests contiennent des questions au niveau d'application et (13 tests) 28,9 % contiennent des questions aux niveaux de connaissance et de compréhension
- La chercheuse a coché « Existe relativement » pour chaque test ne contient quedeux types seulement d'items. De même pour « S'assurer que toutes les questions sont obligatoires »17,8 %des tests contiennent une seule question obligatoire et le reste du test au choix d'un nombre précis de questions.
- Deux matières, phonétiques & lecture et conversation doivent être mesurées à travers des tests de performance orale.

3. **D'après le tableau (1) <u>les normes qui ont dépassé le 80 % sont 10</u> et s'ordonnent comme suit :**

- Laisser d'espaces entre les questions.100%
- Utiliser des illustrations et des dessins de qualité. 100%

- Utiliser des graphiques clairs et compréhensibles (lisibles) .97, 8%
- Eviter les questions en cascade. 97,6%
- Utiliser une langue grammaticalement correcte et sans fautes d'orthographe. 95, 6%
- Citer si le test est en plus d'un page et les numéroter. 93,8%
- La faculté, l'université, la section, la matière et la classe. 93,3%
- Formuler les consignes à l'impératif ou au futur ou à travers une interrogative convenable. 93, 3%
- Numéroter les items en continu. 92,9%
- Utiliser des verbes exprimant des comportements mieux cernés. 84,4%

4. **D'après le tableau (1) <u>il n'y a que deux normes seulement entre le 80% et le 70%</u> et s'ordonnent comme suit :**

- La durée du test. 80%
- S'assurer que toutes les questions sont obligatoires. 71,1%

5. D'après le même tableau ,<u>10 normes ont pris un pourcentage moins de 50%</u>

- La date du test. 11,1%
- Comprendre différents types d'items. (trois types au moins) .11, 1%
- Le nombre des questions à répondre. 20%
- La note indiquée pour chaque question. 31,1%
- Citer les références des textes ou des supports. 33,3%
- Mesurer des objectifs aux niveaux ultérieurs. 35,6 %
- S'assurer que les consignes sont claires et contiennent toutes les données nécessaires. 37,8%
- S'assurer de la signature de l'évaluateur(le professeur de la matière). 37,8%
- La totalité des notes. 48,9%
- Utiliser une phrase exprimant la fin du test.48, 9%

Le tableau (2-A)L'analyse des normes des types de questions.

	Les normes	Existe (2)	Existe relativement (1)	N'existe pas (0)
D	Items de type réarrangement :			
	1. Conserver le début et la fin du texte ou n'importe quelle phrase.	---	---	---
	2. Découper le texte en préservant un minimum de sa prédictibilité.	---	---	---
E	Items à choix multiples :			
	1. Présenter le tronc sous la forme d'une question ou d'un début de phrase.	100	---	---
	2. Ecrire un troncclair et contenu suffisamment de renseignements.	100	---	---
	3. Eviter de donner des détails superflus dans le tronc	100	---	---
	4. Choisir la réponse correcte parmi quatre ou cinq propositions.	50	---	50
	5. Ecrire des distracteurs plausibles, très proches sémantiquement et formellement de la réponse attendue et de la même longueur.	100	---	---
	6. Eviter la forme négative, si on n'a pas le choix insister sur les termes de la négation ou les souligner.	100	---	---
	7. S'assurer qu'aucun indice d'ordre grammatical ou substantif ne risque de révéler la bonne réponse.	100	---	---
	8. Eviter d'utiliser les catégories (toutes les réponses ci-dessus) et (aucune des réponses ci-dessus)	100	---	---
	9. S'assurer qu'il n'y a qu'une seule réponse vraie pour chaque item.	100	---	---
	10. Éviter les énoncés dont la structure est trop complexe.	100	---	---
	11. S'assurer qu'il n'y a qu'une idée par énoncé.	100	---	---
F	Items de type appariement :			
	1. Ecrire plus de réponses possibles que des prémisses.	---	---	---
	2. Elaborer une liste des réponses et des questions pas trop longue (pas plus de 10).	---	---	---
	3. Disposer les questions de façon arbitraire (selon la longueur des mots ou par ordre alphabétique)	---	---	---
	4. S'assurer que la question au complet figure sur une seule page.	---	---	---
	5. S'assurer du caractère homogène des données de chaque question.	---	---	---

Le tableau (2-B) L'analyse des normes des types de questions.

	Les normes	Existe (2)	Existe relativement (1)	N'existe pas (0)
G	Items de type vrai ou faux :			
	1. N'utiliser pas les expressions de la ressource, exprimer les idées dans vos propres mots.	50	50	---
	2. Eviter les phrases dont la structure est trop complexe.	100	---	---
	3. S'assurer que l'item est correct sur le plan syntaxique.0	100	---	---
	4. N'utiliser pas de mots comme (plus, quelques, grand ،quelquefois ou bon) qui ne sont pas suffisamment précis.	50	50	---
	5. S'assurer que toutes les négations sont soulignées ou écrites en caractère gras.	100	---	---
	6. S'assurer qu'il n'y a qu'une idée par item.	100	---	---
	7. Eviter les phrases exprimant des opinions argumentatives.	100	---	---
H	Les textes lacunaires :			
	1. 'assurer qu'ils sont fabriqués à partir de phrases isolées ou un texte.	100	---	---
	2. Ne laisser pas plus de trois trous dans chaque énoncé.	100	---	---
	3. upprimer les mots en sorte que le trou vienne au milieu ou à la fin de l'item.	80	---	20
	4. iter si la longueur du trou reflète la longueur de la réponse attendue ou laisser des lacunes de la même longueur.	60	---	40
	5. ffacer les mots d'un même champ lexical ou d'une même catégorie grammaticale.	40	---	60
	6. S'assurer qu'il n'y a qu'une seule réponse vraie pour chaque trou (lacune).	100	---	---
I	La question combinée (etcompréhension écrite)			
	1. S'assurer qu'elle se compose de 3à 6 questions dont les réponses sont courtes.	84,6	---	15,4
	2. Ordonner les questions de plus facile au plus difficile.	84,6	---	15,4
	3. S'assurer que chaque item mesure une seule tâche.	53,8	38,5	7,7
	4. Mettre chaque question dans une ligne et la numéroter.	84,6	7,7	7,7
J	Items à réponse élaborée (et production écrite)			
	1. Indiquer la longueur du texte à produire. (Le nombre approximatif de mots ou de lignes)	9,7	9,7	80,6
	2. Préciser tous les éléments demandés et/ou indiquer la part de liberté accordée au scripteur.	48,4	51,6	---
	3. S'assurer que chaque item mesure une seule tâche.	83,9	9,7	6,5

Discussion des résultats

1- Items de type réarrangement &Items de type appariement :

- Il n'existe aucun test qui contientdes items de type réarrangement et même de type appariement.

2- Items à choix multiples :

- Deux tests de 45 tests contiennent desitems à choix multiples ça fait 4,4%.
- Un seul test a proposé de choisir la réponse correcte parmi quatre propositions.

3- Items de type vrai ou faux :

- Deux tests de 45 tests contiennentdes items de type vrai ou faux ça fait 4,4%.
- Dans un seul test on a utilisé les expressions de la ressource et on a utilisé des mots qui ne sont pas suffisamment précis.

4- Les textes à trous :

- Cinq tests de 45 tests contiennent des textes à trous ça fait 11,1%.
- Pour la norme « Effacer les mots d'un même champ lexical ou d'une même catégorie grammaticale. », deux tests n'ont pas suivi cette norme car ils mesurent des connaissances apprises par cœur.

5- La question combinée (et compréhension écrite)

- 13 tests de 45 tests contiennent des questions combinées ça fait 28,9%.

6- Items à réponse élaborée (et production écrite)

- 31 tests de 45 tests contiennent desitems à réponse élaborée ça fait 68,9%.
- Indiquer la longueur du texte à produire, la plupart des tests ne précisent pas la longueur du texte ou parfois utilisent des expressions non précises comme un bref, quelques exemples ……..

Enfin, on peut conclure que :

1. Il y a un manque de connaissances et de performances danslesnormes formelles et techniquesde la feuille du test et les normes des types de questions chez les enseignants universitaires.
2. En ce que concernent les normes formelles de la feuille du test,Il n'existe pas un modèle défini de la part de la faculté pour aider les enseignants universitaires à élaborer les tests.
3. Il y a une concentration sur les items à réponse élaborée et la question combinée (qui est à son tour une sorte d'items à réponse élaborée) qui sont faciles à élaborer (de point de vue des enseignants) et en même temps difficiles à corriger si le test ne s'accompagne pas d'un corrigé et d'un barème précis.
4. Les items à correction objective comme ; les textes à trous, les items de type vrai ou faux et les items à choix multiples ,5 tests de 45 tests ont contenu chacun un ou deux types de ces questions et ça fait 11 % seulement des tests. Il est à noter que ces items sont utilisés pour mesurer des objectifs aux niveaux de la mémorisation et de la compréhension.
5. Les enseignants universitaires trouvent de difficulté à élaborer des tests objectifs ou peut être pensent que ce type de tests ne convient pas au niveau universitaire.

Les recommandations

A la lumière des résultats de cette recherche, on recommande de/d' :

1. Entrainer les enseignants universitaires à élaborer les tests à travers des ateliers de travail.
2. Elaborer un modèle de la forme de la feuille d'examen comme guide aux enseignants universitaires.

3. Répandre la culture de l'évaluation et son importance surtout au niveau des tests certificatifs.
4. Etablir une unité à la faculté de pédagogie responsable d'évaluer les tests et donner un feedback aux enseignants universitaires.
5. Elaborer un règlement facultaire d'évaluation des apprentissages convenable à la nature de l'étude à la faculté de pédagogie.
6. Entrainer les enseignants universitaires à former les objectifs des cours et établir l'importance relative des objectifs ce que les aident à mieux choisir les types des questions et mesurer des objectifs aux niveaux ultérieurs.

Recherches proposées

La chercheuse suggère lesrecherches suivantes:

1. Evaluer les tests oraux de la faculté de pédagogie section de françaisà travers des normes universelles.
2. Elaborer un programme d'entraînement pour améliorer les performances des enseignants universitaires à élaborer les tests.
3. Développer les compétences d'auto évaluation chez les enseignants universitaires.

Bibliographie

Références en Français :

1. Aide-mémoire13 : Directives et conseils pour l'élaboration d'épreuves, d'examens écrits, CSFE, Berne, déc2009.http://www.formationprof.ch/download/am/3.pdf(8/1/2012)
2. Beaudet, Suzanne &Monnerie-Goarin, Annie: Petit lexique pour mémoire, LFDLM, n. spécial, août/sept 1993.
3. Brassard, Nancy : Profil de compétences de l'enseignant de niveau universitaire, Université de l'administration publique, Avril 2012, Canada.http://pedagogie.uquebec.ca/portail/doc/profil_nbrassard.pdf.(19/12/2012)

4. Bureau d'évaluation de l'enseignement et des programmes d'études : Evaluation des prestations d'enseignement par les étudiants, Document d'information, Université de Montréal,Novembre2007.http://www.beepe.umontreal.ca/Documents/DocumentDInformationBEPE2007.pdf(15/10/2012)
5. Chaplier, Claire : A l'université, des consignes pour utiliser les consignes, Les Cahiers pédagogiques, n.483, 21 sept. 2010.http://www.cahiers-pedagogiques.com/A-l-universite-des-consignes-pour-utiliser-les-consignes(10/4/2012)
6. Clément, Marlène : Evaluations: formes, moments et acteurs,EDUfle.net11oct.2004http://www.edufle.net/Evaluations- formes-moments-et.html (10/4/2012)
7. Cuq, Jean-Pierre: Dictionnaire de didactique du français langue étrangère et seconde, Paris, CLE international, 2003.
8. Cuq, Jean-Pierre &Gruca, Isabelle : Cours de didactique du français langue étrangère et seconde, Presses Universitaires de Grenoble, 2005.
9. Centre National de Documentation Pédagogique : L'évaluation en Education, Educateur, Ministère de l'éducation algérien,n° 2,juillet, août 2004. http://www.cndp.dz/apps/mourabi/articlerech.php?id=180(20/1/2007)
10. Gilles, Jean-Luc : Feedbacks diagnostiques via internet : Un système individualisé de communication des niveaux de performances des étudiants dans le cadre d'évaluations standardisées, SMART, ULg ,3ème congrès des chercheurs en éducation ,retrouver le plaisir d'enseigner et d'apprendre, 2004http://www.enseignement.be/download.php?do-id=2309&do_check=(13/4/2012)
11. Gilles, Jean-Luc : Construction et gestion qualité d'outils d'évaluation,SMART,Université de Liège, Faculté de psychologie et des sciences de l'éducation , 2001-2002.http://tecfa.unige.ch/tecfa/maltt/actu_recher/seminaireTECFA220306/articles-textes/CGQOE_Introduction.pdf (13/4/2012)
12. Henry, jacques & Cormier, Jocelyne : Impacts des principes didactiques sur l'évaluation sommative. http://discas.ca/Cadreref/documents/intrataxo.html(21/10/2005)
13. Lussier, Denise : Evaluer les apprentissages, Hachette FLE, Paris, 1992.
14. L'université catholique de Louvain : Enseigner à l'université catholique de Louvain, chapitre 3, évaluer les apprentissages, édition 2007

http://www.uclouvain.be/cps/ucl/doc/enseignement/documents/6_Chapitre_III_verrouille.pdf(10/4/2012)

15. Milanovic, M. : Cadre européen commun de référence pour les langues : Apprendre, Enseigner, Evaluer. Evaluation de compétences en langues et conception de tests, Division des politiques linguistiques Strasbourg, octobre2002.

http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Publications/Language_examining_FR.pdf(6/2/2007)

16. Ministère de l'Éducation du Saskatchewan : Evaluation de l'élève, manuel de l'enseignant, cours préliminaire, Janvier 1993.http://www.sasked.gov.sk.ca/docs/francais/tronc/eval/html/a16.htmlhttp://www.sasked.gov.sk.ca/docs/francais/tronc/eval/html/a17.html (9/2/2007)
17. Oasisfle.com : La consigne, chemin qui mène aux apprentissages,http://www.oasisfle.com/documents/consigne_en_apprentissage.htm(10/4/2012)
18. Pendanx, Michèle: Les activités d'apprentissage en classe de langue, Hachette FLE, Paris, 1999.
19. Puren, Christian& autres :Se former en didactique des langues,ellipses,Paris,1998.
20. Quinton, André : Docimologie, Centre de recherches appliquées en méthodes éducatives, Bordeaux2, Mai 2005. http://www.crame.U-bordeaux2.fr/pdf/docimologie.pdf (26/2/2012)
21. Raynal, Françoise &Rieunier, Alain : Pédagogie: Dictionnaire des concepts clés, ESF éditeur, Paris, 1997.
22. Romainville, Marc : L'évaluation des acquis des étudiants dans l'enseignement universitaire, Rapports établis à la demande du Haut Conseil de l'évaluation de l'école, Paris, Décembre2002.http://www.ladocumentationfrancaise.fr/docfra/rapport_telechargement/var/storage/rapports-publics/034000527/0000.pdf (20/10/2012)
23. Salabura, Sébastien : Comment rédiger les consignes pour que les élèves de 6ème soient autonomes lors d'une séance de travaux pratiques, IUFM, Académie de Montpellier, 2002.http://www.crdp-montpellier.fr/ressources/memoires/memoires/2002/b/0/02b0250/02b0250.pdf (10/4/2012)

Références en anglais :

24. Shaaban, Kassim: Assessment of young learners, Forum, vol. 39, n° 4, October 2001.

25. Tenbrink, Terry D.: Assessment, Classroom teaching skills, Houghton Mifflin Company, New York, 1999.

Références en arabe :

٢٦. حلمي الوكيل و حسين بشير:الاتجاهات الحديثة فى تخطيط و تطوير مناهج المرحلة الأولى ، دار الفكر العربي، ١٩٩٩.

٢٧. سوسن شاكر مجيد: أسس بناء الاختبارات و المقاييس النفسية و التربوية، دار ديبونو للنشر و التوزيع،ط١ ،٢٠٠٧.

٢٨. صفاء الأعسر و أخرون : أبعاد التعلم ، تقويم الأداء ، دار قباء ، القاهرة ، ٢٠٠٠

٢٩. على ماهر : مناهج البحث فى العلوم النفسية و التربوية و الاجتماعية ، مكتبة الأنجلو المصرية ، ط٢، ٢٠٠٦/٢٠٠٧

تقويم اختبارات شعبة اللغة الفرنسية بكلية التربية جامعة حلوان في ضوء مقياس لجودة الورقة الإمتحانية

د/ رانيا محمد حامد إسماعيل زيان

- **مستخلص البحث باللغة العربية** :

انطلاقاً من أهمية تقويم الإختبارات الجامعية كنقطة بداية لتحديد الكثير من المشكلات و اتخاذ القرارات المناسبة على مستوى التدريس و التعلم ، فإن هذا البحث يهدف إلي اعدادمقياس لجودة الورقة الإمتحانية المكتوبة باللغة الفرنسية كما يهدف إلى تقويم إختبارات قسم اللغة الفرنسية بكلية التربية جامعة حلوان في ضوء هذا المقياس.

و قد أعدت الباحثة مقياس يتكون من محورين : المعايير الشكلية و الفنية لورقة الإمتحان و معايير أنواع الأسئلة. و تمثلت عينة البحث في ٤٥ إختبارا لشعبة اللغة الفرنسية عن العام الجامعي ٢٠١1/٢٠١2.

و قد ابرزت نتائج البحث وجود قصور لدى أعضاء هيئة التدريس في معارف وأداءات المعايير الشكلية و الفنية لورقة الإمتحان و معايير أنواع الأسئلة إلى جانب التركيز بشكل كبير علي الأسئلة المقالية و المركبة.

****